



Vers une détection et une classification non-supervisées des changements inter-images

Alexandre Fournier, Xavier Descombes, Josiane Zerubia

► To cite this version:

Alexandre Fournier, Xavier Descombes, Josiane Zerubia. Vers une détection et une classification non-supervisées des changements inter-images. Traitement et Analyse de l'Information : Méthodes et Applications (TAIMA), May 2007, Hammamet, Tunisia. 2007. <inria-00271158>

HAL Id: inria-00271158

<https://hal.inria.fr/inria-00271158>

Submitted on 8 Apr 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers une détection et une classification non-supervisées des changements inter-images

Alexandre Fournier, Xavier Descombes et Josiane Zerubia

Projet Ariana (INRIA/I3S)
2004, route des Lucioles, B.P. 93
06902 Sophia Antipolis Cedex France
`prenom.nom@sophia.inria.fr`
`http://www-sop.inria.fr/ariana/`

Résumé L'objectif de ce travail est la détection et la classification de changements sur des couples d'images satellitaires panchromatiques haute résolution en vue de la détection de cibles spécifiques.

La démarche adoptée est composée de deux parties :

- dans la première partie, nous effectuons une analyse en composante principale itérative afin d'obtenir une probabilité de changement au niveau pixélique. Nous effectuons ensuite une régularisation par champ de Markov sur les données obtenues afin d'avoir une décision « dure » tout en incluant des informations de connexité dans le processus.
- dans la seconde partie, nous nous attaquons au problème de segmentation non supervisée des pixels qui ont été retenus comme « modifiés » au cours de l'étape précédente. La technique utilisée est une adaptation de l'algorithme des *k-moyennes*. Celle-ci prend en compte un terme entropique permettant de déterminer le nombre optimal de classes pour la segmentation.

Mots clés changements, détection, classification, EM, champs de Markov, entropie, k-moyennes.

Remerciement Les auteurs remercient la Délégation Générale pour l'Armement (DGA) pour la fourniture des couples d'images Quickbird. Le premier auteur remercie également la DGA et le Centre National de la Recherche Scientifique (CNRS) pour le soutien financier partiel de sa thèse .

1 Introduction

La détection de cibles sur des images satellitaires haute résolution peut être d'une importance primordiale dans un contexte militaire ou civil. On s'attache ici, en particulier, au problème de la détection de cibles en considérant les changements induits entre deux images de la même zone géographique prises à des dates différentes. Si l'utilisation de deux images au lieu d'une seule peut apporter des informations, elle génère également une grande quantité de fausses alarmes.

- en premier lieu, les erreurs de recalage peuvent avoir une influence considérable sur le taux de fausses alarmes [8], et la détection de changement doit donc être suffisamment robuste par rapport à ces erreurs ;
- par ailleurs, en supposant un recalage parfait entre les deux images considérées, d'autres changements peuvent être détectés qui ne peuvent être considérés comme résultant de l'apparition ou de la disparition de cibles potentielles. Parmi les phénomènes générant les fausses alarmes les plus fréquentes en zone urbaine, on citera notamment :

- les fausses alarmes dues aux différences générales d’ensoleillement ou à des différences d’appareil de prise de vue ;
- les fausses alarmes dues aux effets de stéréoscopie ;
- les fausses alarmes dues aux changements d’ombres portées dues à des différences d’orientation lumineuse.

2 Discrimination des pixels changés/inchangés

Afin de s’affranchir des fausses alarmes dues à une différence générale d’illumination, on cherche à déterminer une caractéristique de « non-changement » entre les images au temps T_1 et T_2 (*i.e.* déterminer une transformation linéaire « optimale » entre la radiométrie des deux images). Si l’on considère le graphe de l’intensité lumineuse des pixels au temps T_2 en fonction de l’intensité lumineuse au temps T_1 , alors cette caractéristique se présente comme un nuage de points allongé. Les pixels changés ont tendance à s’éloigner de ce nuage. On aurait tendance à privilégier une approche reposant sur une Analyse en Composante Principale (ACP). Cependant, cette approche peut se révéler biaisée : en effet, la composante principale est le vecteur \mathbf{w} qui maximise la variance de la projection d’une distribution sur son axe :

$$\mathbf{w}_1 = \arg \max_{\|\mathbf{w}\|=1} E \left\{ (\mathbf{w}^T \mathbf{x})^2 \right\} \quad (1)$$

En conséquence, cette approche est particulièrement sensible à la présence de points situés à grande distance du nuage de points que l’on cherche à caractériser. Pour pallier ce problème, les auteurs de [7] proposent une démarche itérative en deux étapes :

1. pondérer chaque point de l’histogramme en fonction de la distance qui le sépare de la composante principale ;
2. recalculer la nouvelle composante principale en fonction des points pondérés.

Cette méthode donne de bons résultats, mais ne tient pas compte de la notion de connexité. Dans le cadre de la détection de changements, cette information est capitale pour pouvoir replacer un pixel dans son contexte. L’algorithme précédent rend, par ailleurs, une probabilité de changement p_{Ch} pour chaque pixel P , alors que l’on cherche à lui assigner une étiquette « changé » (Ch) ou « non changé » (Nc). On effectue alors sur cette carte de probabilités une segmentation se fondant sur les champs de Markov. On définit V , le terme d’attache aux données par pixel en fonction de p_{Ch} .

$$VD_{\text{Nc}}(P) = \frac{1}{1 - p_{\text{Ch}}(P)} \quad (2)$$

$$VD_{\text{Ch}}(P) = \frac{1}{p_{\text{Ch}}(P)} \quad (3)$$

La notion de connexité est introduite par le biais du terme de probabilité *a priori* qui se fonde sur le modèle d'Ising [1] :

$$VA(i) = -\beta \sum_{\{P, P'\}, P \sim P'} (1 - \delta(\omega(P'), \omega(P))) \text{ où } \begin{cases} \omega(P) \in \{\text{Ch}, \text{Nc}\} \\ \sim \text{ désigne la relation de voisinage} \\ \delta \text{ désigne le symbole de Kronecker} \\ \beta \in \mathbb{R}^{+*} \end{cases} \quad (4)$$

On optimise l'énergie totale en utilisant un algorithme de recuit simulé [3,4] .

3 Classification par algorithme des k-moyennes

Une fois la détection des changements effectuée, il reste à discriminer parmi les changements les cibles potentielles. Une première étape consiste alors en une classification des changements en différentes catégories. Les méthodes d'*espérance-maximisation* (EM) semblent indiquées pour effectuer cette classification, assurant une convergence assez rapide. Néanmoins, les algorithmes tels que l'algorithme classique des k-moyennes demande la connaissance *a priori* du nombre de classes. [6] propose une approche permettant de s'affranchir de cette connaissance en ajoutant un terme d'entropie dans l'énergie à minimiser.

3.1 Le modèle utilisé

Soit $X = \{\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N\}$ l'ensemble des N couples de pixels qui ont été classés comme « changés » au cours de l'étape précédente. On cherche à déterminer une partition \mathcal{C} de l'espace des valeurs de ces couples $\Lambda = \lambda_{T_0} \times \lambda_{T_1}$ (avec λ_{T_0} et λ_{T_1} les niveaux de gris des images respectivement prises au temps T_0 et T_1). On remarquera qu'il s'agit des niveaux de gris définis par les valeurs effectivement prises par les pixels du couple d'image, et non pas les valeurs pouvant être prises (pratiquement, une image codée sur 8 bits peut donc engendrer un λ ne contenant pas certains entiers de $[0, 255]$). On définit donc $\mathcal{C} = \{c_i \text{ avec } i \in [1..k]\}$ où $\forall (i, j) \in [1..k]^2, i \neq j \Rightarrow c_i \cap c_j = \emptyset$ et $\bigcup_{i=1}^k c_i = \Lambda$.

L'algorithme classique des *k-moyennes* propose de classifier l'information en cherchant à minimiser le terme énergétique suivant :

$$U = \sum_{i=1}^k \sum_{j \in c_i} \frac{T(\mathbf{x}_j - \mathbf{y}_i) \Sigma^{-1}(\mathbf{x}_j - \mathbf{y}_i)}{2} \quad (5)$$

Où \mathbf{y}_i représente la moyenne des couples de points situés dans la classe c_i et Σ représente la matrice de covariance de l'ensemble des couples de points de X . Le nombre de classes k est alors un des paramètres de l'algorithme. On remarque que la diminution du nombre de classes induit un minimum d'énergie plus important et la segmentation ainsi obtenue risque d'être peu représentative de l'information portée par la sous-image X . *A contrario* un nombre élevé de classes permet une meilleure minimisation du critère énergétique et une possibilité d'énergie U nulle pour $k = N$ et chaque classe contenant un couple de niveaux de gris unique ; la sur-segmentation obtenue alors n'est évidemment porteuse d'aucune analyse

de l'information. Le choix du nombre de classes est, en conséquence, critique vis-à-vis des résultats d'un tel algorithme et doit être lié à l'information intrinsèque portée par l'image. Un des critères permettant d'appréhender à cette information est l'entropie statistique.

3.2 Terme d'entropie

Pour un espace partitionné en k classes, on peut définir l'entropie du système comme :

$$\mathcal{S} = - \sum_{i=1}^k p_i \log(p_i) \quad (6)$$

où $p_i = \frac{\text{Card}(c_i)}{N}$ est la probabilité d'avoir un couple de pixels dans la classe c_i .

On remarque que l'entropie est minimale avec un partitionnement pour lequel une classe est prédominante et les autres classes sont de mesure nulle.

On cherche ici à trouver un compromis entre ce terme d'entropie et l'énergie que l'on cherche à minimiser avec l'algorithme classique des *k-moyennes*. La nouvelle forme de l'énergie à minimiser est alors :

$$U = \sum_{i=1}^k \sum_{j \in c_i} \frac{T(\mathbf{x}_j - \mathbf{y}_i) \Sigma^{-1}(\mathbf{x}_j - \mathbf{y}_i)}{2} - \alpha_0 \sum_{i=1}^k p_i \log(p_i) \quad (7)$$

On reconnaît dans le terme de gauche, l'énergie à minimiser dans l'algorithme classique des *k-moyennes*. On cherche, par ailleurs, à ce que l'influence du terme d'entropie soit indépendante de la taille de l'image. Or, le terme énergétique est lui directement proportionnel à N . Le terme α_0 s'écrit donc sous la forme $\alpha_0 = \alpha \cdot N$. On peut alors reformuler l'équation (7) sous la forme :

$$U = \sum_{i=1}^k \left[\sum_{j \in c_i} \frac{T(\mathbf{x}_j - \mathbf{y}_i) \Sigma^{-1}(\mathbf{x}_j - \mathbf{y}_i)}{2} - \alpha \cdot \text{Card}(c_i) \log(p_i) \right] \quad (8)$$

$$\text{soit : } U = \sum_{i=1}^k \sum_{j \in c_i} \left[\frac{T(\mathbf{x}_j - \mathbf{y}_i) \Sigma^{-1}(\mathbf{x}_j - \mathbf{y}_i)}{2} - \alpha \cdot \log(p_i) \right] \quad (9)$$

L'algorithme de *k-moyennes* entropique adapté à la détection de changements peut alors s'écrire de la manière suivante :

1. on initialise les *K moyennes* pour les classes avec $K \gg k_{\sim}$ où k_{\sim} est une estimation grossière du nombre de classes que doit comporter la partition. Dans le cas où Λ est faible, on peut choisir d'initialiser les moyennes sur les éléments de Λ , mais dans le cas général, cette solution s'avère trop coûteuse en temps de calcul. On choisira d'initialiser les moyennes des classes selon une distribution gaussienne dont les moyennes et la matrice de covariance sont les estimations de celles de X , ou simplement d'initialiser les moyennes sur des couples pris au hasard dans X [2].

2. **espérance** on assigne chaque couple \mathbf{x}_j à la classe la plus proche au sens de la nouvelle distance définie plus haut :

$$\forall \mathbf{x}_j \in X, \quad x_j := c_{i_0} : i_0 = \arg \min_i \frac{T(\mathbf{x}_j - \mathbf{y}_i) \Sigma^{-1}(\mathbf{x}_j - \mathbf{y}_i)}{2} - \alpha \cdot \log(p_i) \quad (10)$$

maximisation pour chaque classe, on recalcule la moyenne et la probabilité d'appartenance d'un couple de pixels :

$$\forall i \in [1..k], \text{Card}(c_i) \neq 0, \quad \mathbf{y}_j := \sum_{j \in c_i} \frac{\mathbf{x}_j}{\text{Card}(c_i)} \quad (11)$$

$$\forall i \in [1..k], \quad p_i := \frac{\text{Card}(c_i)}{N} \quad (12)$$

test si les \mathbf{y}_j ont été modifiés revenir à l'**espérance**, sinon, fin de l'algorithme

4 Résultats

Les couples d'images utilisées représentent la ville d'Abidjan, en Côte d'Ivoire, et ont été prises à 6 mois d'intervalle. Les effets de stéréoscopie et de changement d'ombres portées sont nettement visibles à travers la présence de bâtiments à forte élévation dans la partie centrale de l'image.



Fig. 1. Couple d'images *Quickbird* originales, © distributeur SPOT Images, fournies par la DGA

L'algorithme d'Analyse en Composante Principale et la régularisation par champ de Markov permet d'obtenir un masque de pixels changés sur lequel on peut effectuer la partie ultérieure de l'analyse.

Enfin, les résultats de la classification finale sont indiqués sur la figure 2.

5 Conclusion

Nous avons présenté une technique de détection et de classification des changements entre deux images au niveau pixelique. Les résultats sont également régularisés à l'aide d'un

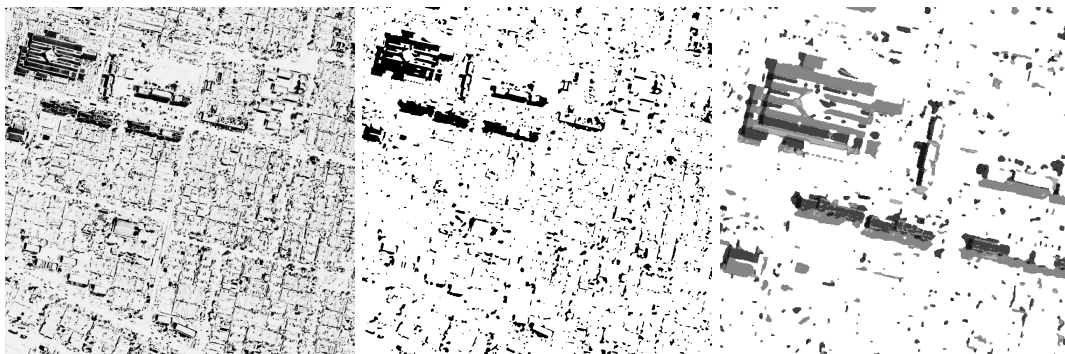


Fig. 2. Probabilité de changement par pixel (à gauche), régularisation par champ de Markov (au centre) et classification issue de l'algorithme de k -moyennes entropique (à droite). La régularisation élimine notamment les fausses alarmes dues à des pixels isolés. Afin d'éviter d'avoir une classe trop prépondérante, nous avons adopté un schéma de décroissance géométrique pour le paramètre α [5]. Le nombre final de classes est de 10. On remarque que les ombres portées sont divisées en deux classes différentes, selon leur image d'appartenance.

champ de Markov. Néanmoins, une partie des informations se situe à un niveau supérieur aux niveaux pixélique et local. Nous travaillons donc actuellement sur une approche de type *régions* afin de reconnaître certains motifs récurrents (notamment au niveau des bâtiments et de leurs ombres portées), et de tenir compte des informations de taille et de forme des zones.

Références

1. Julian Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2) :192–236, 1974.
2. Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, août 2006.
3. S. Geman et D. Geman. Stochastic Relaxation, Gibbs Distributions and Bayesian Restoration of Images. *IEEE trans. PAMI*, 6 :721–741, 1984.
4. P. J. M. Laarhoven et E. H. L. Aarts, éditeurs. *Simulated annealing : theory and applications*. Kluwer Academic Publishers, Norwell, MA, Etats Unis, 1987.
5. A. Lorette. *Analyse de Texture par Méthodes Markoviennes et par Morphologie Mathématique : Application à l'Analyse des Zones Urbaines sur des Images Satellitales*. Thèse de doctorat, Université de Nice Sophia Antipolis, septembre 1999.
6. G. Palubinskas, X. Descombes, et F. Kruggel. An unsupervised clustering method using the entropy minimization. Dans *Proc. International Conference on Pattern Recognition (ICPR)*, Brisbane, Australie, août 1998.
7. R. Wiemker. An iterative spectral-spatial Bayesian labeling approach for unsupervised robust change detection on remotely sensed multispectral imagery. Dans *Proc. Computer analysis of images and patterns*, Kiel, Allemagne, septembre 1997.
8. D. Xiaolong et S. Khorram. The effects of image misregistration on the accuracy of remotely sensed change detection. Dans *IEEE Trans. Geosci. Remote Sensing*, volume 36, pages 1566–1577, septembre 1998.